Oracle NoSQL Database

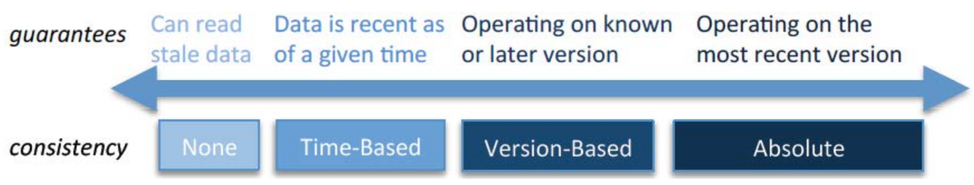
Siddhardha N, Shuai Wang
Department of Computer Science and Engineering
University of Bridgeport, Bridgeport, CT

Abstract

Data is stored and retrieved from Database in tabular format which is called Relational Database and familiar from 1960's as SQL (Structured Query Language) Database. Today with the rapidly increasing collected information, data driven applications are rising in science and business territories. From 21st century the term NoSQL is being popular and it is also called 'Non SQL' or 'Non-Relational' or 'Not only SQL'. For large scale applications on datacenters or cloud distributed NoSQL systems are well known for their ease of use. There are many NoSQL non-relational Databases are introduced depending on its type based on the CAP theorem. NoSQL database data models are classified into different categories like Key- Value system, Document based system, column based system and Graph based system mainly. In this paper we will be introducing Oracle NoSQL database which is horizontally scaled with high availability, key valued database for cloud and web services. It has load balancing which is transparent even after dynamically adding new capacity.

Technical Overview

The Oracle NoSQL database which is inherently a scale out type of database which is met to run across lots of independent pieces of hardware instead of scaling up on master piece of hardware. In addition to that it is key value database it supports ACID transactions at the same time it also embraces BASE which is known as Basically Available Soft state Eventually consistent like other NoSQL databases.

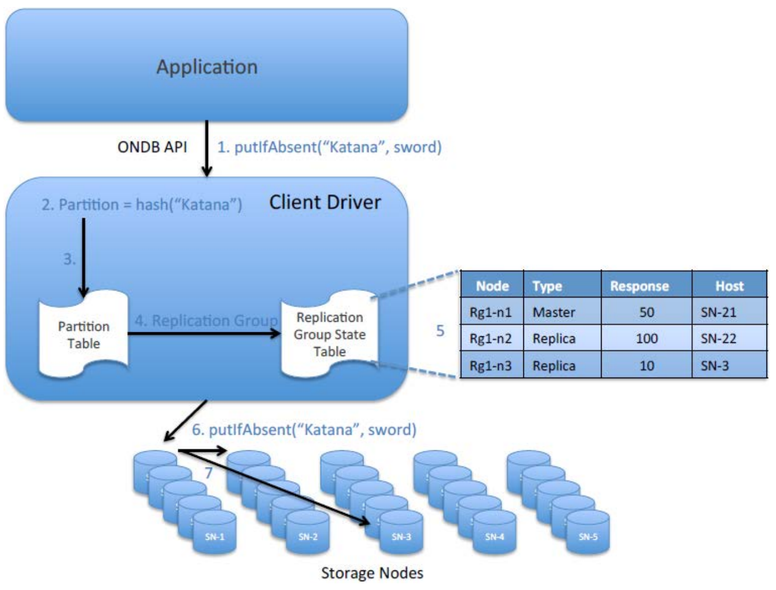


It is highly available and pretty simple to administer. The technologies using by Oracle NoSQL database are dealing with web scale applications like Amazon, real time event processing, mobile data management, sensor data capture, online gaming and many other different places. It is not like traditional database which we used in the past. Above figure indicates how the scope of flexible consistency strategies enables engineers to effectively make business arrangements giving information ensures while meeting application latency and scalability requirements. This NoSQL is used in Hadoop environment where Big data come into place. Parallel processing is done using this NoSQL database on big data. When performing CURD operations we will have Exceptions like Consistency Exception, Request time out Exception, Fault Exception in this database. This technology integrates with ETL processing i.e. Extract, Transform and Load which makes it fast, reliable and distributed storage to applications.

Architecture

Oracle NoSQL Database architecture presented by fallowing the execution operation in the upcoming diagram. In this we can observe an application which is linked to NoSQL client driver. That is library which you link to your application. The driver is connected to store, inside it a bunch of data partitions and they are isolated and they don't share anything. All of the data is separate from each of those. i.e. basically data is hashed in different partitions . Each of this partition is responsible for replicating the data for some number of times for reliability purposes.

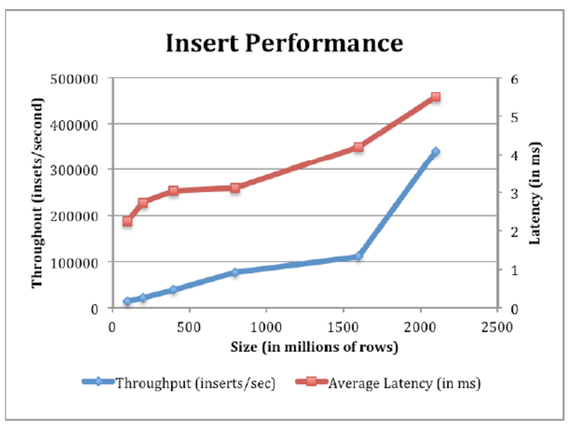
These partitions are elastic in nature, we can add more of them, split them and contract them. Write operations are directed to particular master node which is elected at that time to handle all write operations. And from that it is replicated to all other nodes handling availability.

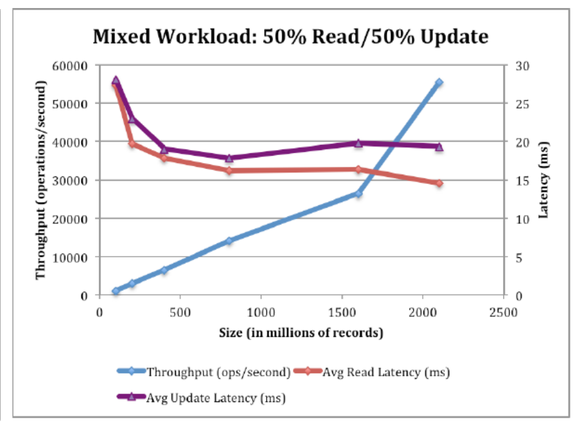


At the same time read operations can be done from any of those nodes in entire store. Because of this it will try to make more partitions so that we can have more read operations through out storage node, overall read through put will reduce the latency. In Physical architecture those partitions mentioned are racks. If one master node read the data it will replicate to other nodes in different partitions for High availability. And these are isolated hardware zones. The driver will monitor all of the process in the node and keep a track of it. We can also isolate those partitions in geographically different regions.

Performance

First graph is the performance about Oracle NoSQL Database raw insert, which shows a solitary replication group. It has 3 hubs and stores 100,000,000 records. The system has 32 mirroring groups with 96 hubs, it stores 2.1 billion records. The result shows that the throughput in blue and the reaction time in red. We can tell that the throughput have a linearly relationship with database size and number of replication groups develops and only have a little bit increment in response time.





The right graph shows the performance of half read workload and half updates, the throughput and response time. When the system has a huge data size and replication groups, the read and update latency will decreases while straightly scaled throughput.

Conclusion

Oracle NoSQL Database is a distributed key-value database for a group between nodes can be configured as a storage system provides highly reliable, scalable and available data storage. Data is stored as key-value pairs, according to the hash value of the primary key will be written to a specific storage node. With a copy of the storage node, which is fast failover to ensure high availability, a node failure, and the best query load balancing.